
Creating and Capturing Artificial Emotions in Autonomous Robots and Software Agents

Claus Hoffmann^{1,*}, Pascal Linden² and Maria-Esther Vidal³

¹*Research Group Robots and Software Agents with Emotions, Sankt Augustin, Germany*

²*University of Bonn, Germany*

³*TIB Leibniz Information Centre for Science and Technology, Hannover Germany*
E-mail: hoffmann.claus@web.de; s6palind@uni-bonn.de; Maria.Vidal@tib.eu

**Corresponding Author*

Received 30 October 2020; Accepted 03 May 2021;
Publication 24 June 2021

Abstract

This paper presents ARTEMIS, a control system for autonomous robots or software agents. ARTEMIS can create human-like artificial emotions during interactions with their environment. We describe the underlying mechanisms for this. The control system also captures its past artificial emotions. A specific interpretation of a knowledge graph, called an Agent Knowledge Graph, stores these artificial emotions. ARTEMIS then utilizes current and stored emotions to adapt decision making and planning processes. As proof of concept, we realize a concrete software agent based on the ARTEMIS control system. This software agent acts as a user assistant and executes their orders and instructions. The environment of this user assistant consists of several other autonomous agents that offer their services. The execution of a user's orders requires interactions of the user assistant with these autonomous service agents. These interactions lead to the creation of artificial emotions within the user assistant. The first experiments show that it is possible to

Journal of Web Engineering, Vol. 20_4, 933–970.

doi: 10.13052/jwe1540-9589.2043

© 2021 River Publishers

realize an autonomous user assistant with plausible artificial emotions with ARTEMIS and record these artificial emotions in its Agent Knowledge Graph. The results also show that captured emotions support successful planning and decision making in complex dynamic environments. The user assistant with emotions surpasses an emotionless version of the user assistant.

Keywords: Autonomous agents, artificial emotions, agent knowledge graphs.

1 Introduction

The development of autonomous agents able to act and decide in complex environments is one of the essential early visions of Artificial Intelligence (AI). Such agents should manage unforeseen problem situations and emulate human behavior to resolve these problems. (compare [36]) Despite many advances (e.g., in multi-agent systems and semantic technologies that improve the interaction between agents), this vision has only partially become a reality. Specifically, developing autonomous agents capable of *adapting* to complex environments is not yet completely solved. In such environments, it is vital to have the ability to plan and re-plan if necessary (compare [46]) and have the capacity to modulate the execution of the plan by adaptive decision-making. As Dörner and Güss argue, “The function of emotions is to adjust behavior considering the current situation” [12, p. 308]. This idea’s foundation is the observation that humans can often adapt to complex and dynamic environments quite successfully. Thus, an idea to transfer this ability from humans to agents is to equip autonomous agents with *artificial emotions*.

Problem Statement We address the following two problems: (i) how to create artificial emotions that are a realistic simulation of human emotions in certain problem situations and (ii) how to capture these artificial emotions in a knowledge graph of an autonomous agent.

Proposed Solution We propose ARTEMIS, a robot and software agent control system, which enables agents to create, capture, and utilize artificial emotions. ARTEMIS exploits the generated and stored artificial emotions to modulate behavior and decision making. The PSI theory of the cognitive psychologist Dörner [12] provides the theoretical foundation for ARTEMIS’s

general structure. Additionally, the Component Process Model (CPM) of the emotion psychologist Scherer [39] provides the theoretical basis for essential aspects of the creation of artificial emotions in ARTEMIS. Thus, ARTEMIS and its artificial emotions rely on a solid theoretical background, briefly introduced in Sections 4.1 and 5.1. Knowledge bases are essential components of autonomous robots or software agents. They are the cornerstone for their planning and decision-making. There are several ways to realize such a knowledge base. We suggest for this purpose a particular version and interpretation of knowledge graphs. Our interpretation of knowledge graphs, we call Agent Knowledge Graph, is intended to support autonomous robots and software agents in planning and decision making in complex environments by storing artificial emotions.

Our Contributions. We present the design of our robot or software agent control system ARTEMIS. The control system is capable of creating and capturing artificial emotions. The basis for creating artificial emotions is the appraisal of interactions of the agent with other autonomous agents. Both cognitive processes and need processes are involved in realizing these appraisals. We demonstrate how ARTEMIS implements both types of processes. The ARTEMIS control system contains an Agent Knowledge Graph, which stores the emotions and makes them available for later planning and decision-making processes. This paper is an extension of an earlier paper by us [19]. We enrich the previously presented concepts with further details. First, give a more extensive overview of state of the art. Furthermore, we present a revised representation of the architecture of ARTEMIS. We describe Dörner's PSI theory in more detail and elaborate on its connection to ARTEMIS. We also go more in-depth about Scherer's appraisal patterns and show how ARTEMIS realizes them. We also provide a more comprehensive specification about how ARTEMIS generates and stores emotions.

The paper is structured as follows. Section 2 motivates a possible application area of ARTEMIS. Section 3 discusses related approaches and their relevance to the ARTEMIS control system for autonomous agents. In Section 4, we look at Dörner's PSI theory as the foundation of ARTEMIS. Then we discuss the general architecture of ARTEMIS. In Section 5, we discuss how ARTEMIS creates emotions. In Section 6, we devise an Agent Knowledge Graph to model the motivating example's problem. In Section 7, we present our experimental study and describe our experimental results. In Section 8, we discuss our conclusions and our future work.

2 Motivating Example

We motivate our approach using a typical situation that may be present in a wide variety of data-driven scenarios. Examples of application scenarios include selecting (a) machines in future ‘Smart Factories’, (b) means of transport in ‘Supply Chains 4.0’, and (c) information sources by an autonomous information broker in a ‘decentral dataspace’ like an ‘Industrial Data Space.’ Our exemplary application scenario moves within the context of the so-called Service Web (see [13]). With this exemplary scenario, we can study principal problems of service selection without getting lost in the details of concrete application areas. The process of the exemplary application scenario is as follows (see Figure 1). An autonomous agent takes on the role of an autonomous user assistant for its user. The autonomous user assistant accepts the orders of its human user. To execute an order, the autonomous user assistant searches its knowledge base for a suitable plan. A plan defines a list of steps. For each plan step, the autonomous user assistant must find a suitable service agent that performs the step. Autonomous service agents offer their services at different prices and are differently trustworthy. The autonomous user assistant has to decide which service agent fits best with the

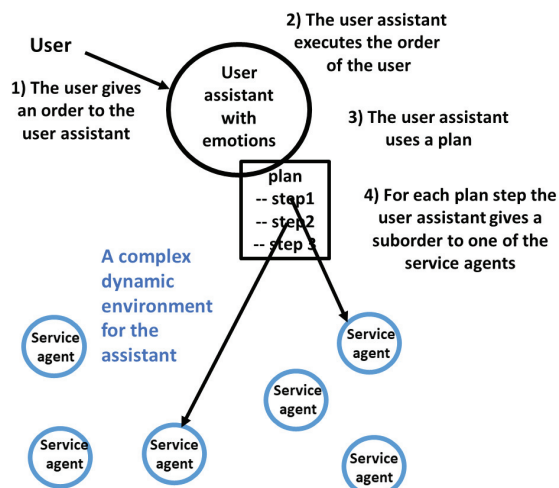


Figure 1 Motivating Example. An autonomous user assistant executes a user’s order in a complex environment. For this purpose, it uses a plan. The user assistant selects the most suitable service agents to execute the individual plan steps. The service agents are also autonomous. Several interactions take place between the user assistant and the service agents. Appraisals of these interactions by the user assistant create artificial emotions.

current situation. The following conditions form the basis for the exemplary application scenario:

Condition 1. In complex dynamic environments (e.g., ‘Industry 4.0’ applications), conditions for cooperation with autonomous service agents can change from time to time. Present cooperation partners may leave the environment of the autonomous user assistant, and others may arrive. As a result, the search for suitable cooperation partners becomes a permanent task.

Condition 2. The cooperation partners of the autonomous user assistant are autonomous themselves and try to maximize their outcomes. Therefore, the results of cooperation are often uncertain. It is always possible that the user assistant’s cooperation partners do not meet the agreements and provide results that do not fulfill its expectations. Several situations may cause this violation of the user assistant expectations. One reason could be that cooperation partners are not capable of delivering their promised services. Another reason could be that they did not understand the mandate correctly. Lastly, they may deliberately did not execute the job correctly to gain an advantage.

These conditions provide the basis for a complex interaction between the autonomous user assistant and the autonomous service agents. Appraisals of these interactions create corresponding emotions in the user assistant. For example, ‘Excited’ when something goes well in contrast to expectations (and the result was significant) and ‘Disdainful’ when a cooperation partner performs poorly (and it is possible to balance this out). Through numerous interactions with the service agents, the user assistant gains experience on cooperation partners’ reliability. Emotions are created and stored in the Agent Knowledge Graph of the user assistant. With these emotions, the user assistant gains essential knowledge overtime to help future effective planning and decision-making. Conventional approaches without artificial emotions would only determine whether an interaction was successful or not. The emotion-based approach, on the other hand, is much more differentiated. Emotions summarize the agent’s assessment of the entire underlying situation. An essential function of emotions, which we utilize in ARTEMIS, is to adapt the planning and decision-making of an autonomous actor to a particular situation. Scherer [39] describes this as follows: “Emotions are mechanisms that enable the individual to adapt to constantly and complexly changing environmental conditions” (from [39]). This fact applies to both current and remembered emotions.

3 Related Work

Research in computer science and emotions currently focuses on recognizing user emotions. Other approaches try to recognize emotions in texts, human faces, or the language (see [18]). This research direction has already achieved significant results. Emotion analysis will be essential for machines to react appropriately to their human users' emotions. Such analyses are, therefore, crucial for the next step in human-computer interaction (HCI). However, the approach presented in this paper is not about the recognition of human emotions. Instead, the focus is on creating and memorizing artificial emotions in autonomous agents. We use these artificial emotions to adapt autonomous robots and software agents' behavior to the respective environment. It is also crucial that the presentation of these artificial emotions (e.g., face, voice, or gestures) can help users understand the system's decisions and actions. The basis of this understanding is that human users can often imagine that they probably would have had similar emotions in similar situations and that they would have acted or decided similarly on this foundation. The approach ARTEMIS presented in this paper has two results. On the one hand, it serves to improve the performance of autonomous agents. On the other hand, it is also a contribution to the research area of HCI.

The basis for generating artificial emotions in agents is research in psychology, cognitive architectures of cognition research, and agent architectures of artificial intelligence. We briefly describe the essential foundations of these three research areas. After that, we present a selection of approaches based on these foundations and discuss their relation to our approach.

3.1 Psychological Emotion Research and Autonomous Agents with Artificial Emotions

So-called appraisal theories are currently predominant in psychological emotion research. These theories describe emotions as judgments within cognitivism and emphasize the close connection between cognitive and emotional processes. Appraisal theories deal with appraisals of those characteristics of events of the environment that are important for an organism [31]. These appraisals generate emotions and, in this way, corresponding adaptation reactions. Well-known appraisal theories include the theory from Smith and Lazarus [41] Ortony, Clore, and Collins [33], and Scherer [39].

Smith and Lazarus. Smith and Lazarus [41] made essential contributions to stress research and emotion research. They regarded emotions as evolutionary

strategies that aim primarily at eliminating a threat to an organism's motives. Smith and Lazarus distinguished between a primary appraisal, a secondary appraisal, and a reappraisal. A primary appraisal is assessing the significance of occurred or future events for the motives of an organism. The assessment of an organism's possibilities for action and coping is the subject of a secondary appraisal. Reappraisal means a re-evaluation of the situation, which can lead to a modification of the primary assessment. Smith and Lazarus were mainly engaged in stress research.

OCC. The most widely used model for agents with artificial emotions is the emotion model from Ortony, Clore, and Collins [33], or the OCC model for short. The OCC model describes the characteristics of prototypical situations and associates these situations with emotions. The OCC model defines 22 different emotions and relies on variables that describe their intensity. OCC appraisals are hierarchical so that they form a decision tree. An important reason for the OCC model's popularity is that its fundamentals suggest a computer implementation [20]. That means this model is quite accessible to computer scientists. The OCC model describes evaluations at a very high level of abstraction. The evaluations rely on a purely cognitive assessment of the consequences of events, agents' actions, and aspects of objects. It does not refer to needs or motivational processes.

Scherer. Another critical appraisal theory is the Component Process Model (CPM) from Scherer [39]. Within his model, Scherer describes the components involved in the emotional process, their interaction, and the effects of different appraisal results. An important detail is that Scherer emphasizes that realistic emotions require cognitive processes and need processes. (compare [39]) We use a part of Scherer's CPM model in ARTEMIS. In Section 5.1 of this paper, we discuss this separately. To us, the theory of Scherer seems to be the most promising to realize systems that show believable emotions and use emotions to modulate the behavior of autonomous agents. For this reason, for the creation of artificial emotions, we focus our work on Scherer's theory.

3.2 Classification of Emotion Models

Emotions psychologists distinguish between discrete and dimensional emotion models. We summarized the first contributions in both areas.

Basic Emotions. Several scientists assume that there exist so-called basic emotions. One of the most prominent representatives of this idea is Paul Ekman [14]. An essential aspect of basic emotions is that they are discrete,

which means each emotion is fundamentally different from every other basic emotion. They represent different essential and often occurring (crisis) situations and the reactions that have become evolutionarily accepted as the best solution. (compare [14]) Basic emotions are innate, are present in all people, and are triggered automatically. Emotions psychologists assume that there are at least 6, but possibly up to 15 basic emotions in humans. ARTEMIS uses not a model of basic emotions but a dimensional emotion model.

Dimensional Models of Emotions. The assumption with dimensional models is that a few independent dimensions can characterize emotions. Points or regions in the space spanned by these dimensions define emotions. There are many different dimensional models. Differences exist in the number of dimensions, the dimensions' names, and the emotions' coordinates. In ARTEMIS, we rely on the 3-dimensional model of Mehrabian [29, 30]. This model uses the dimensions Pleasure (P), Arousal (A), and Dominance (D). Therefore, it is commonly called the PAD model.

3.3 Cognitive Architectures and Autonomous Agents with Artificial Emotions

An important question is how to realize the cognitive processes of autonomous agents. The best known traditional cognitive theories are Soar (States, Operators, And Results) [24] and ACT-R* (Adaptive Control of Thought – Rational) [1]. However, Soar and ACT-R* are pure cognitive theories. These theories model human behavior as problem-solving with relation to a given task. However, these two theories do not model needs or even refer to concepts like motivation or emotion. Soar and ACTR* can be classified as computer science approaches rather than psychological approaches. (compare Detje [7, p. 86]) Dörner's PSI theory, however, is a general theory of autonomous systems. It describes the interaction between need processes, cognitive processes, and emotional processes. It describes, from a psychological perspective, how motives and intentions arise from needs. Scherer emphasizes in his CPM theory that "The existence of needs and values in agents are, of course, the central prerequisite of a computational agent model. Without needs or goals, no real emotions." (from Scherer [39, p. 87]). The Dörner's PSI theory represents a solid framework to realize autonomous agents with artificial emotions; it provides the foundations for ARTEMIS.

3.4 Traditional Architectures for Autonomous Agents From the Area of Artificial Intelligence

There are four different traditional agent architectures: logic-based, reactive, hybrid, and BDI agents (compare [21]).

Logic based agents. Logic-based agents operate in worlds which are subject to restrictions. For example, the world cannot change while the agent decides what to do. (compare [21]) A logic-based agent architecture is therefore not suitable for complex dynamic environments.

Reactive agents. A reactive agent selects actions based on the current sensor information and a set of condition action rules. Local information on its environment is the basis of the behavior of a reactive agent. Reactive agents are relatively easy to implement. Unfortunately, the application areas of such agents are quite limited. (compare [21])

Hybride agents. These architectures are combinations of different agent types. Robotics research often uses such agent architectures. (compare [21])

BDI agents. The three main components of the BDI architecture for agents are Belief, Desires, and Intentions. The philosopher, Michael Bratman at Stanford University, developed 1987 the so-called practical reasoning system. According to Bratman, the system of practical reasoning should reflect the process of human practical thinking. On this basis, Anand Rao and Michael Georgeff developed in 1991 the BDI architecture for autonomous agents [34]. BDI agents know their environment (beliefs), desirable states (desires), and currently pursued intentions. BDI (the currently most crucial traditional architecture) uses the terms desires and intentions but does not use motivational processes such as a need system as a basis for that. Also, the BDI architecture has a philosophical and not a psychological origin. In contrast, both Scherer's CPM theory [39] and Dörner's PSI theory [9] are psychological theories. Scherer emphasized the importance of a motivational system with needs to create artificial emotions. An essential part of Dörner's PSI Theory is a model of needs, from which motivations and intentions arise.

3.5 A Selection of Already Realized Approaches for Agents with Artificial Emotions

There are diverse approaches to create agents with artificial emotions. We only discuss a selection of these (additional references [25, 28], and [23]).

One line of this research is how to show emotions to improve human-agent interaction. Examples for this can be found in [2, 5], and [4]. Besides aspects of communication, some approaches also deal with the influence of emotions on agent acting and decision-making. In the following, we focus on some selected approaches and present the main differences in the creation of emotions between these approaches and our ARTEMIS approach.

Crucial – in both the PSI theory and ARTEMIS – is that intentions are grounded in motivations and get their meaning through these motivations. Furthermore, motivations are grounded in needs processes and get their meaning through these needs processes. The definition of this requires an appropriate psychological theory. This theory must describe the connections and interactions between needs, motives, and intentions. In this way, motivations and intentions can get meaning.

We will briefly discuss several existing approaches for agents with emotions. None of these approaches models needs or motives based on a well-founded psychological theory that describes the interaction of need, motivation, and intention processes. The presented approaches use terms like need, motivation, or intention purely intuitively or heuristically. Owing to this reason, they do not fulfill Scherer's requirement for the creation of emotions.

Abbots. Cañamero [6], based the creation of artificial emotions on a motivational system. The system used drives, motivations, and emotions to select behaviors. Behavior aims to satisfy needs related to the motives. Although the approach of Cañamero describes important concepts of agent decision making, the approach is missing a grounding of the concepts of needs and a psychological theoretical basis for the interplay of emotional, need-based, and cognitive processes.

Cathexis. Velásquez [43] is one of the first researchers who dealt with an emotion-based decision-making system for autonomous robots. Velásquez used emotions to activate different behaviors, to generate attention, to create appropriate emotional expressions, and to make it possible to learn from past experiences. Velásquez uses drives, motivations, and emotions to select behaviors. For the realization of autonomous agents with emotions, Velásquez made many valuable proposals. However, with his approach, the problem is that a grounding of drives, motivations, and emotions is missing.

Kismet. Breazeal [5] developed a sociable robot named Kismet. The robot has a simple model of needs, where the needs represent the specialized three goals of Kismet: engaging with people, engaging with toys, and resting from time to time. Three dimensions (arousal, valence, and stance) describe the

emotions of Kismet. Self-crafted appraisals can generate values for these three dimensions to activate one of nine implemented artificial emotions. The active emotion then influences the robots posture, facial and vocal expression. It can also trigger certain emotion specific behaviors. Problematic is that Breazeal uses the concept of the drive purely intuitively as a mechanism to switch between the three main goals implemented in Kismet. Therefore, the need system and the interplay with motivational and emotional influences have no theoretical foundation.

The ALEC agent architecture. Gadanho [15] presents artificial emotions that can influence behavior. Her approach uses drives, motivations, and emotions to select behaviors. Emotions provide input to a reinforcement learning algorithm. The enforcement system takes the emotion with the highest intensity for reinforcement. The approach uses only cognitive appraisal processes for that purpose.

EDBI. In 2007 Jiang et al. [21] presented EBDI, an architecture for emotional agents. The goal of the project was to add an emotional component to the BDI architecture. BDI is the basis of EBDI. BDI does not model needs, and so does not EBDI. That means there is no grounding of the emotions used in the approach.

The EMA-Model. Marsella and Gratch [27] presented the EMA-Model. EMA stands for Emotion and Adaption. The SOAR architecture [24] is the basis of EMA on the cognitive side. On the emotional side, the EMA-Model basis is the appraisal theory of Smith and Lazarus [41]. SOAR is a purely cognitive approach. It does not model needs or motivations.

SOAR-Emote.Model. Marinier et al. [26] presented the SOAR-Emote.Model. SOAR realizes the cognitive part of the SOAR-Emote.Model. However, the authors stress that they could just as well have used ACTR*. As emotion theory, the authors use Scherer's appraisal theory. Furthermore, the authors use Newell's theory of cognitive control PEACTION [32] to model the agent's decision-making behavior. SOAR is a purely cognitive approach. It does not model needs or motivations.

An emotion-based decision-making approach. Antos and Pfeffer [3] formulate emotions as mathematical operators that update the agent's goals relative priorities. In their approach, artificial emotions can also influence the importance of goals. This approach presents many excellent ideas. For example, agents should adapt to the current situation by acting more cautiously if there is some danger or acting more bravely if everything seems to run well.

However, the authors base their system on purely cognitive appraisals of the situation. The approach does not model grounded intentions, motivations, and needs. The authors use these concepts, only heuristically / intuitively.

Emotions and their Use on a Decision-Making System. The approach of Salichs and Malfaz [37] deals with the generation and role of artificial emotions in the decision-making process of autonomous agents. The basis of the approach is drives (needs), motivations, and artificial emotions. The authors regard artificial emotions as fear, happiness, and sadness. In this approach, agents can appraise the environment and raise an artificial emotion due to this appraisal. A pure cognitive appraisal is a basis for the evaluation of the environment.

The WASABI Architecture. Becker-Asano [4] presented the WASABI architecture for believable agents. WASABI incorporates many different approaches related to emotions. These are the PAD model, the OCC model, and Ekman's basic emotions. The controlling emotion system has three components – emotions, moods, and boredom. For most of the emotional concepts, Becker-Asano uses Ekman's basic emotions. WASABI does not model grounded intentions, motivations, and needs. The author uses these concepts, only heuristically/intuitively.

4 The Foundations of ARTEMIS

The basis for essential parts of ARTEMIS is the PSI theory of the cognitive scientist Dörner [12]. The PSI theory formalizes cognitive, motivational, and emotional processes and their interaction. The theory defines a computational architecture for autonomous agents. Dörner implemented virtual agents based on his PSI theory and showed that they could adapt to a simulated, complex environment and learn how to reach their own goals within this environment. We will first provide an outline of Dörner's PSI theory and then present the architecture of our approach ARTEMIS.

4.1 Preliminaries: An Outline of Dörner's PSI Theory

The PSI theory ([9, 10, 12]) defines an architecture for autonomous systems (in the following called PSI agents). The origin of a PSI agent's behavior regulation is its motivational processes. The basis for these processes is a homeostatic system evolving around five basic human needs: existential needs (energy, water, pain avoidance), a need for sexuality, a social need

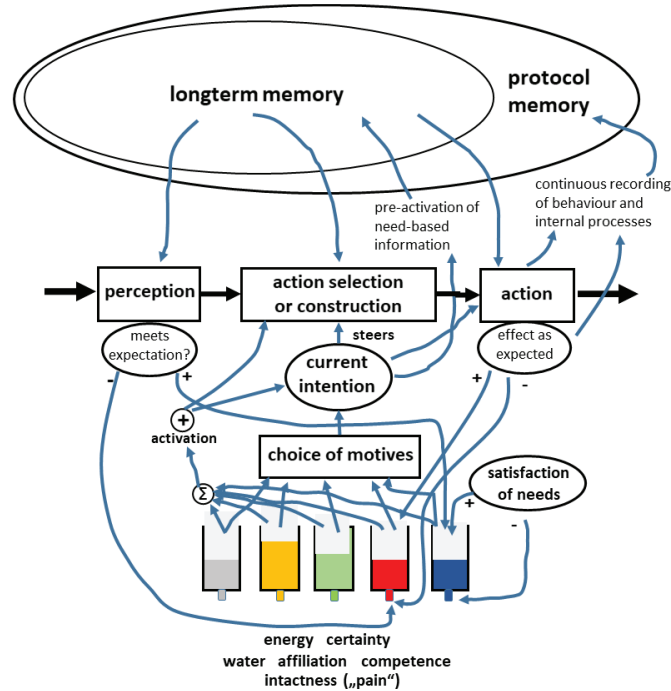


Figure 2 PSI. Dörner’s PSI theory is the basis for the ARTEMIS architecture (see Figure 3) and realizing Scherer’s appraisal pattern (see Figure 4). This figure shows an overview of the structure of PSI (cut-out and own translation from [9]). The PSI theory defines an architecture of autonomous agents. This Figure primarily describes the interaction of need processes, motivations, intentions, and cognitive processes.

for affiliation, a cognitive need for certainty, and a second cognitive need for competence. A set of “tanks” – one for each need (see Figure 2) – can represent this need system. A tank contains a varying amount of “liquid” that empties over time or by experiencing negative events. When the filling level falls below a certain threshold called setpoint, a demand arises, which is stronger, the greater the derivation from the setpoint is. This demand is the motivational signal for the PSI agent to take action.

When a demand signal appears, the PSI agent creates a motive. A motive is a structure containing references to the current situation, the associated need, and a goal to satisfy the present demand. The PSI agent can have multiple motives at once, but a selection process only chooses one motive to become dominant as the agent’s current intention. The basis for this selection is the strength of the indicated needs and the subjectively estimated

likelihoods to satisfy them. The prerequisite for this is the current situation of the PSI agent and its knowledge from previous experiences.

For the current intention, the PSI agent initiates a cognitive planning process. That means the PSI agent tries to construct an action sequence that leads from the current situation to the goal associated with the current intention. To do that, the agent draws on his experiences. If none of its existing knowledge can be applied, a trial-and-error approach is used, creating new knowledge. This new knowledge might become the basis for future planning if some of the executed actions lead to the need's satisfaction or mitigating the need.

A PSI agent aims to generate pleasure signals when demand is partially or completely satisfied by an action or event. A pleasure signal increases the filling level of the respective need tank again. These pleasure and displeasure signals also reinforce the paths leading to positive, satisfactory, or negative, need-increasing events as part of a PSI agent's learning process. In the PSI theory, learning means that the strength of connections in the PSI agent's memory is changed, or new connections are created. The memory of a PSI agent consists of a particular neural network structure Dörner invented (for more details on that, see [12]).

PSI models emotional processes not as distinct entities that become active on certain events. Emotions instead serve as a continuous modulation system, which is highly interrelated with the other processes. Three modulation parameters realize emotions – activation, resolution level, and selection threshold. These parameters are derived from the need system, representing the PSI agent's current internal state. All of the PSI agents' core processes as perception, accessing memory, planning, and executing actions are adjusted according to the three modulation parameters' current values. This solution allows the PSI agent to adapt its behavior to the current demands dictated by the environment and based on its internal state. In the PSI theory, "emotions" are considered identical to these modulation parameters' current configuration and the resulting adjusted cognitive processes.

4.2 The Architecture of ARTEMIS

ARTEMIS's core is closely related to Dörner's PSI architecture of autonomous systems, especially the processing of needs, selecting motives, and handling the current intention (see Section 4.1.). However, there are several own solutions within ARTEMIS, which we discuss in the following.

Memory. An Agent Knowledge Graph realizes ARTEMIS's memory, not as a particular form of neural networks, as the PSI theory realizes it. The

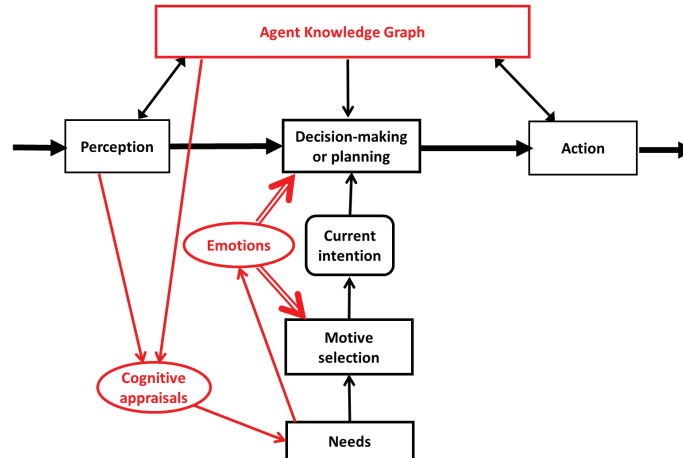


Figure 3 The Architecture of ARTEMIS. Dörner’s PSI theory provides the basis for this architecture’s essential components. Some new components are specific to the ARTEMIS architecture. In ARTEMIS, an Agent Knowledge Graph realizes the memory of Dörner’s PSI. In contrast to Dörner’s approach, ARTEMIS has a specific cognitive appraisal component closely connected to ARTEMIS’s need system. The need system and the appraisal component produce values for the parameters ‘Pleasure’, ‘Arousal,’ and ‘Dominance,’ which define artificial emotions in the PAD cube. These artificial emotions, or better expressed their Pleasure, Arousal, Dominance components, influence the ‘decision-making/planning’ and ‘goal selection’ components of the control system. The Agent Knowledge Graph captures these artificial emotions.

reason for this is that we do application-oriented research in the field of artificial intelligence. PSI’s approach to realizing a memory by a particular form of neural networks is highly exciting. However, more intensive and more prolonged basic research is needed before this approach will be ready for practical application. Therefore, in our opinion, a preliminary realization based on knowledge graphs seems to be appropriate for the near future.

Appraisals and emotions. In PSI, the realization of emotions relies on the three needs generated parameters “resolution level”, “selection threshold”, and “activation” (see [12]). The values of these three parameters lead to emotions and modulation of cognitive processes. However, PSI does not address appraisal processes explicitly. ARTEMIS models appraisal through explicitly described cognitive assessment processes that work closely with needs. The results of the appraisals act as a very detailed evaluation of the current situation under different aspects. In ARTEMIS, this also generates efficiency and inefficiency signals corresponding to the PSI theory’s

pleasure and displeasure signals. These signals increase or decrease the filling levels of the need tanks. The needs then create the values of the three parameters Pleasure, Arousal, and Dominance. We discuss ARTEMIS needs and appraisals further in Section 5.2. The parameters, Pleasure, Arousal, and Dominance, are used in ARTEMIS to define artificial emotions and modulate the autonomous agent's cognitive and behavioral processes. Dörner discussed in his book [10] that the three parameters "Lust-Unlust (pleasure-displeasure)," "Erregung-Beruhigung (excitement-calming)," and "Spannung-Lösung (stress-solution)" mentioned by Wilhelm Wundt (compare [10, p. 220]) as components of emotions would also fit well to his theory. Wundt's parameters are a precursor of the PAD parameters developed later by Mehrabian [29, 30] and can be easily mapped to them. Dörner did not pursue this idea; however, he recently assured the authors of this paper – in a personal conversation – the feasibility of this approach.

Modulation. The three parameters, "resolution level," "selection threshold," and "activation," which are used in PSI to define emotions, are also generated by ARTEMIS. As in PSI, ARTEMIS uses these parameters to modulate cognitive processes. However, ARTEMIS additionally uses the ARTEMIS-specific PAD parameters for modulation. For the definition of artificial emotions in ARTEMIS, we rely entirely on the three PAD parameters because scientifically based mappings of emotions exist (see [29, 30]).

5 Creating Artificial Emotions

We focus on describing how ARTEMIS creates artificial emotions. The basis for creating artificial emotions is the Component Process Model (CPM) theory of the emotion researcher Scherer [39]. The CPM theory defines emotions as a result of evaluating external events, objects, or situations by an agent. This article shows that ARTEMIS can realize the theoretical evaluation scheme defined by the Scherer's CPM theory, and the use of appraisals as a basis for artificial emotion creation. We present Scherer's CPM theory next.

5.1 Preliminary: Scherer's Component Process Model (CPM)

Scherer provides a comprehensive appraisal theory with his Component Process Model (CPM) developed in 1984 [38]. Within the framework of his model, he describes the various components necessary for creating emotions in detail. He explains how these components interact with each other. Scherer

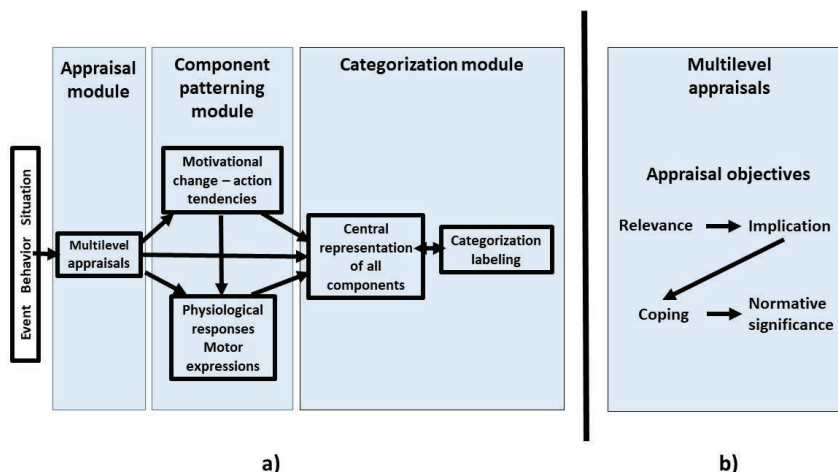


Figure 4 Architecture of Scherer’s Component Process Model of Emotion Scherer [39, p. 49]. (a) The Component Process Model (CPM) consists of five subsystems that realize five different functions. These functions are (1) Appraisal of objects, events, or situations (Cognitive component), (2) preparation and direction for action, (3) signaling of behavioral intention, (4) Central representation of all components, (5) assignment to fuzzy emotion categories and labeling with emotion words. (b) The most critical component in CPM is the appraisal module. Within the appraisal module, the multilevel appraisal component realizes four appraisal objectives. These appraisal objectives are divided into individual appraisal criteria, each of which has its check defined. Own figure based on [39, p. 50]).

emphasizes at multiple points that his CPM could be particularly suitable for computer-aided modeling. However, Scherer also mentions the requirements and sometimes significant challenges involved if a computer-based agent should implement his evaluation criteria. The Component Process Model (CPM) consists of five subsystems that realize five different functions. These functions are: (compare [39, p. 49])

- **Cognitive component:** Evaluation of objects and events.
- **Motivational component:** Preparation and direction for action.
- **Motor expression component:** Signaling of behavioral intention.
- **Neurophysiological component:** Regulation of internal Subsystems.
- **Subjective feeling component:** Monitoring of internal state and external environment.

On Outline of Scherer’s Appraisal Pattern for Events. The appraisal module is the most basic and most important in the overall architecture (compare [39, p. 50]). “An event happens, and we will instantly examine its relevance

by drawing from memory and motivation and attend to it immediately if it is considered relevant.” (from Scherer [39, p. 47])

Scherer describes the four primary appraisal goals that an appraisal process of events must realize to enable an organism to react adaptively to an outstanding event. Evaluation goal 1: How relevant is this event? Does it directly affect the organism or its social reference group? (relevance). Evaluation goal 2: What are the implications or consequences of this event? How do they affect the well-being and immediate or long-term goals of the organism? (implications). Evaluation goal 3: How well can the organism cope with or adjust to these consequences? (coping potential). Evaluation goal 4: What is the significance of this event for the organism’s self-concept and social norms and values? (normative significance) (compare Scherer [39, p. 50]).

Scherer subdivides the four appraisal objectives into more detailed appraisal criteria. The appraisal criteria include novelty, valence, goal relevance, urgency, goal congruence, responsible agent, coping potential, and norms (see Scherer [39, p. 51]). With his proposal, Scherer presents a theoretically sound appraisal pattern. However, he does not give any precise information on how to realize it. However, Scherer gives hints on the boundary conditions for implementation; he also emphasizes the existence of needs and goals as essential prerequisites for the appraisals of events. Further, Scherer’s criteria point out that computational agents who have no needs or goals cannot have real emotions (Scherer [39, p. 52]).

5.2 Creating Artificial Emotions in ARTEMIS

We expand the knowledge of events that have taken place to include knowledge about the artificial emotions associated with them. ARTEMIS creates emotions which are the result of the agent’s appraisals of events (see Figure 4). First, we further discuss ARTEMIS’s need system, which is directly influenced by the appraisal results. This need system then generates values for the parameters ‘Pleasure’, ‘Arousal,’ and ‘Dominance.’ The parameter values generated by the need system are then mapped to the PAD cube of emotions and define emotions there (Figure 5). After these foundations, we discuss how ARTEMIS realizes the appraisal pattern defined in Scherer’s theory.

5.2.1 The need system of ARTEMIS

Here, we present seven needs captured in the ARTEMIS control system. Why does our control system work with these seven needs as opposed to PSI? The answer is: Dörner uses the needs shown in Figure 2 in the context of his psychological research questions. Dörner defines needs for the following

scenario: “A robot must survive on an island. It needs water and energy and has the task to collect something”. The motivating example contains an entirely different scenario. It describes service agents’ assignment to perform steps of a user assistant’s plan (see the motivating scenario). We have adapted the user assistant’s needs to this scenario because different needs are necessary depending on this scenario.

Instead of needs like “water” and “energy,” the user assistant has the need to “preserve existence.” The need for sexuality has been modified in ARTEMIS as a need to “be agile.” The other needs proposed in the PSI theory (pain avoidance, affiliation, certainty, competence) are also used in ARTEMIS. Additionally, we implemented a need to “avoid damages,” which extends the need to “avoid pain” so that the agent also avoids to cause pain or damage to its environment. As a result, the following requirements are satisfied by ARTEMIS.

1. **Preserve existence.** That is, being able to execute orders, and make sure that services can be paid.
2. **Avoid pain.** For robots, it could mean to avoid structural damages. For software agents, it could mean not spending too much money.
3. **Be agile.** Change methods and maybe partners from time to time, also they should neither get bored nor boring.
4. **Affiliation.** The need for robust social integration and a good relationship with others.
5. **Certainty.** Being knowledgeable about the environment. Certainty results from the ability to explain and predict events based on knowledge about the environment.
6. **Competence.** Being capable of effectively and efficiently dealing with real-world problems.
7. **Avoid damages.** For robots, it means maintaining machines or buildings and not overloading machines. For software agents, it represents the ability of not making decisions that endanger the environment.

These needs are closely related to ARTEMIS’s emotional model. As a means of internal representation, we use a dimensional theory of emotion, namely the PAD cube, to characterize emotions. To derive values for the three dimensions “Pleasure,” “Arousal,” and “Dominance,” we use the ARTEMIS need system in the following way:

- **Pleasure.** Rising and falling of the strength of needs determine the level of pleasure.
- **Arousal.** A combination of the strengths of all needs determines the level of arousal.

- **Dominance.** The levels in the tank of the need for certainty and the need for competence determine the agent's dominance.

The needs' strength depends on the corresponding levels (represented with the variable L) of the associated need tanks. ARTEMIS calculates the levels of the need tanks continuously. The level can only take values between 0 and 1. Let eff , $ineff$, $Cert$, $Comp$ be efficiency, inefficiency, certainty, and competence, respectively. Sig indicates the current amount of efficiency or inefficiency signals, which are mainly generated from the appraisals of events. The efficiency and inefficiency signals have a weight of W_S , which models the strength of their impacts. The needs also have a specific weight of W_N , representing their general priority. (see [11])

With these definitions, we give a schematic overview of how ARTEMIS adjusted needs and derives values for Pleasure, Arousal, and Dominance.

```

Pleasure := 0
Arousal := 0
for (i := 1 to NumberNeeds)
  L[Need[i]] :=  $W_S(eff) * Sig(eff) - W_S(ineff) * Sig(ineff)$ 
  L[Need[i]] := MAX(0, MIN(1, L[Need[i]]))
  Need[i] :=  $\ln(1 + L[Need[i]])$ 
  Pleasure += L[Need[i]]
  Arousal += Need[i] *  $W_N(i)$ 
Dominance := Need[Cert] * (1 - Need[Comp])

```

5.2.2 Realizing Scherer's appraisal pattern in ARTEMIS

ARTEMIS realizes the appraisal steps that Scherer defined in his CPM theory. The appraisals defined by Scherer often involve interactions between cognitive processes and need-based processes. Therefore, alongside the necessary cognitive calculations, the need system also plays an essential role in the ARTEMIS appraisal system (see Figure 3). The results of these processes together then form the actual appraisals. The dynamics of the need processes involved generate values of the PAD parameters (as described above) and thus the emotions associated with these parameters. Scherer's CPM theory describes the information processing in humans that leads to emotions. All in all, his theory describes very complicated interrelations, and as Scherer himself stresses, it would be challenging to realize for artificial agents in all its aspects. However, because our goal is to build artificial autonomous agents, we first tackle partial areas of the problem. The motivating example provides the context for such a sub-area. The motivating example helps make

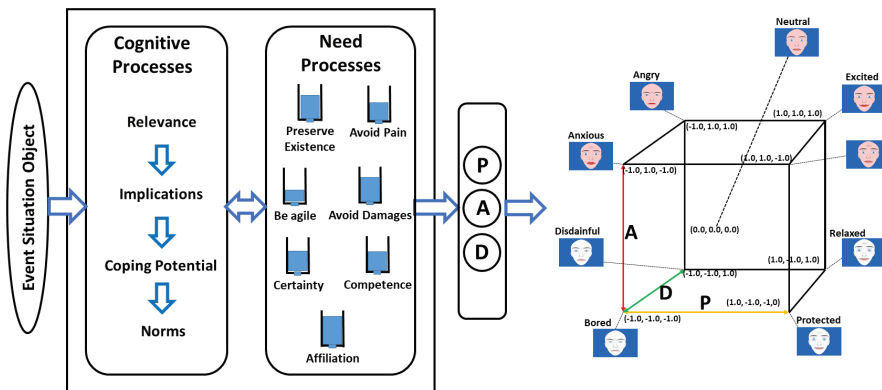


Figure 5 Generating artificial emotions in ARTEMIS. Scherer’s appraisal pattern [39] defines appraisal objectives and appraisal criteria for the evaluation of external events. ARTEMIS uses the results of these appraisals as inputs for the need system which represents the agent’s internal state, we derive values for pleasure, arousal, and dominance, which characterize the agent’s current emotional state in the PAD cube of emotions.

the matter manageable; thus, fundamental questions are analyzed here first. Not all needs are used for the appraisal steps presented here. Other needs such as “Avoid pain” or “Be agile” are essential for the modulation of “Motiv selection” or “Decision making or planning.” Modulation is not the topic of this paper, but we will present these processes and how the need system is involved in this modulation in a forthcoming paper. In the following, we describe how we realized Scherer’s appraisal patterns in ARTEMIS. We present Scherer’s description of the respective appraisals first. Secondly, we explain how ARTEMIS can perform the appraisals.

Appraisal Objective 1: Relevance

The first three tests check the relevance of an event, i.e., they check whether it is worthwhile for the agent to continue to deal with the matter. Such relevance detectors mainly prepare for further tests.

- A check for *novelty*
 Each new stimulus requires attention and further checks. It may involve a potential danger or unexpected benefit. (compare Scherer [39, p. 51])
 The solution in ARTEMIS. The ARTEMIS system evaluates the novelty of an event compared to the agent’s expectations. The PSI theory calls an agent’s expectations regarding the development of the current situation, “expectation horizon.” Dörner and Stäudel [8] describe this as follows:

“The expectation horizon is formed based on the given situation and “reality models.” If x is the case and if I know that x is a concrete form of X and if usually, X turns into Y , then y will be the case.” This kind of conclusion, in a straightforward way, creates the horizon of expectation. Thus, the horizon of expectation represents an extrapolation of the given situation based on the knowledge of its usual evolution. Own translation from (from Dörner and Stäudel [8, p. 309]). The better the agent’s reality model is, the fewer unexpected events for the agent will occur. This first check for novelty is only a relatively superficial analysis. Detailed analyses are carried out in the “check for the probable outcome” and “check of failure to meet expectations.”

- A check for *intrinsic pleasantness*

This check evaluates whether a perceived stimulus is more likely to result in pleasure or pain for the system – independently of its current internal state. (compare Scherer [39, p. 51])

The solution in ARTEMIS. The basis for this check by ARTEMIS is the stored emotions in the Agent Knowledge Graph. Since in this paper, we are moving in the motivating example, we deal with emotions associated with service agents. ARTEMIS represents pleasure and pain by stored positive or negative emotions in the Agent Knowledge Graph.

- A check for *relevance to goals and needs*

The system checks how important a particular event is for one’s own needs or goals. An event can touch for a single or for several needs or goals at the same time. Different needs and goals can, in turn, have different importance. (compare Scherer [39, p. 52])

The solution in ARTEMIS. This check identifies whether and if so, which needs (general) or objectives (specific) the event concerns. This check selects affected needs for further checks. Needs (and related objectives) have different importance in ARTEMIS. For example, the need to “Avoid pain” has a higher priority than the need to “Be agile.”

Appraisal Objective 2: Implications

The second appraisal objective deals with the potential consequences of an event and how the agent is affected by them. (compare Scherer [39, p. 52])

- A check for *cause*

Essential information about an event is its cause. Two crucial parts of this information are about agency and intention. One question is: who did it? Another question is: why did he or she do it? (compare Scherer [39, p. 52])

The solution in ARTEMIS. In the context of the motivating example (Section 2), only interactions between the user assistant and service agents are possible. Therefore, the “agency” question is easy to identify: the answer follows from the interaction protocol. What is interesting here is an assessment of the intentions. For example, suppose a service agent fails to deliver. In that case, there is a possibility that the service agent is attempting to defraud the user agent or that an external problem has occurred for which the service agent is not responsible. How can the user agent distinguish between these possibilities? In ARTEMIS, we have solved this in the context of the motivating example. In the case of a service agent associated with positive emotion, the user agent first assumes that it is not responsible for the result. However, if the service agent is related to a negative emotion, the user agent is more likely to believe the service agent intended the wrong result. No need is directly affected by this check.

- A check for *probable outcome*
This check determines “the probability of inevitable consequences of the event.” (compare Scherer [39, p. 52])
The solution in ARTEMIS. In this check, the system evaluates how likely a deviation from the outcome associated with the event is with the help of its expectation horizon. ARTEMIS associates this assessment for a probable outcome with the need for “Certainty.” An event that is not within the range of expectations shows the agent’s internal system that its reality model is flawed. Thus the agent’s need for certainty increases.
- A check for *failure to meet expectations*
In this check, the system calculates the probability, whether the situation caused by the event is consistent with the agent’s expectations or not. (compare Scherer [39, p. 53])
The solution in ARTEMIS. This check of failure to meet expectations works similarly to the check for novelty, but provides a more in-depth and more detailed analysis. As this check is also closely related to the check for probable outcome, it is associated to the need for certainty in the same way.
- A check for *conduciveness to goals and needs*
The more an event contributes directly or indirectly to achieving objectives, the higher is its need or goal conduciveness. Conversely, the more an event disturbs or even blocks the satisfaction of a need or the achievement of a goal (compare Scherer [39, p. 53])

The solution in ARTEMIS. In the context of the motivating example (section 2), this check is relatively easy to realize. If a service agent communicates successful progress or good results, the user agent considers that as conducive. Any result associated with delays or bad results is evaluated as obstructive. Primarily involved is the need “Preserve Existence,” but the need “Certainty” is also affected. If a result is available earlier than expected, this will reduce the need for “Preserve Existence.” If the availability of a result is unexpectedly delayed, this increases the strength of the need. For the need “Certainty,” it is essential whether the user agent expected the result or not. Suppose the negative result is entirely unexpected, the need for “Certainty” increases.

- A check for *urgency*

When an event threatens high priority needs or goals, an organism’s appropriate response is urgently needed, because a delayed response could potentially aggravate the situation. (compare Scherer [39, p. 53])

The solution in ARTEMIS. The needs in ARTEMIS have different weights and ARTEMIS has the necessary information to calculate the quantities of time course and contingency at all times. These two components are combined with the need strength to evaluate this check for urgency. The need to “Preserve Existence” is affected. If an event is considered very urgent, this will increase the level of this need.

Appraisal objective 3: Coping potential

Organisms are not required to endure the effects of events that affect them passively. By taking appropriate action, they may avoid or at least mitigate the effects of adverse events. To explore the possibilities for action for an agent, Scherer defines the check for control, power, and coping potential. (compare Scherer [39, p. 54])

- A check for *control*

Control is the extent to which any natural agent can influence or control an event. (compare Scherer [39, p. 54])

The solution in ARTEMIS. This check is purely cognitive. Can the result be influenced at all, or is it something like a natural phenomenon? Or better, in the context of the motivating example, is it possible that the user assistant can change anything? For example, is there even the necessary time to change something in the result? This check influences the need for competence.

- A check for *power*

Power is the ability of the agent to change results according to its interests. (compare Scherer [39, p. 54])

The solution in ARTEMIS. This check is purely cognitive. If there is the possibility for control, which resources are available to the user assistant to avoid the result. For example, is there still enough money and time to realize an alternative solution? This check influences the strength of the need “preserve existence” and the need “competence”.

- A check for *potential for adjustment*

The adaptive potential is the ability of an agent to adapt to the effects of an event if they cannot be changed. (compare Scherer [39, p. 54])

The solution in ARTEMIS. This check is purely cognitive. For example, ARTEMIS checks if it is possible to delay the deadline for the end result without severe consequences and how fast the next best alternative can be achieved. This check influences the strength of the need “preserve existence” and the need “competence”.

Appraisal objective 4: Normative significance

The question is, what an event means compared to a reference group’s norms and compared to the agent’s norms, values, and self-concept. (compare Scherer [39, p. 55]) In ARTEMIS’s current state, we define the desirable norms with which the agent should evaluate events and actions. When someone respects or violates these norms, the positive or negative emotions affect the agent’s relationship with the interaction partner. As the need for affiliation changes, the values of the PAD parameters change as well.

- A check with *external standards*

This check for external standards evaluates to what extent an action is compatible with the standards of a reference group. (compare Scherer [39, p. 55])

The solution in ARTEMIS. The norms to which ARTEMIS refers in the context of the motivating example are the principles of the so-called “honorable merchant” (German: “Ehrbarer Kaufmann”): “The Honourable Merchant behaves honestly in business dealings with customers and suppliers. Keeping promises, especially concerning the state of the art and safety as well as quality, delivery time, and payment terms are the basis for relationships with customers and suppliers”. (Own translation from [44]. For a version in English, see [45]).

If someone does not play by the rules, this triggers the negative emotions of the user assistant. The reverse is also true, but the negative emotions are more potent than the positive. ARTEMIS searches its memory to see if there have been similar incidents in the past. If this is the case, it reinforces the negative evaluation of the event. The system directly links the negative emotion to the responsible service agent (i.e., it weakens an

existing positive emotion; it strengthens an existing negative one). This influences the need for affiliation.

- A check with *internal standards*

This check evaluates to what extent an action is compatible with the own norms and values. (compare Scherer [39, p. 55])

The solution in ARTEMIS. The basis for internal standards is currently the same principles of the “honorable merchant” as external principles. Possibly we will change this in later implementations. The involved need is the same as with the check with external standards.

5.2.3 Mapping the PAD parameters to artificial emotions

The PAD parameters form a cube, as shown in Figure 5. The values of these parameters correspond to different points in this cube. There are different proposals for mappings the points or regions of the PAD cube to emotions in the literature. For our approach, we lean on the emotion mapping from Mehrabian [29,30]. Mehrabian considers only octants (subcubes) of the PAD cube. However, it makes perfect sense to name the extreme points of the PAD cube after these octants. So-called dimensional approaches make it possible to define vague boundaries of emotion categories. In our approach, the intensity of the eight emotions associated with the octants increases from the center to the edges of the cube. An essential aspect of our approach is, that the artificial emotions created by ARTEMIS are not arbitrary character strings, but they have a meaning or grounding. In fact, there are several layers of meaning associated with the artificial emotions created by ARTEMIS. These meanings can be derived as follows: An emotion in ARTEMIS is represented by its pleasure, arousal, and dominance values in the PAD cube. According to Gaerdenfors [16] the PAD cube already has the potential to equip emotions with meaning. In our approach, the PAD values of an emotion are derived from the agent’s internal state represented by the need processes. The need processes themselves are influenced by the appraisals of events. That makes an emotion in ARTEMIS an abstract and meaningful representation of the agent’s evaluation of events with respect to its internal state.

6 Capturing Artificial Emotions – The Agent Knowledge Graph

Previously, we described the process followed by ARTEMIS to create artificial emotions. In this section, we explain how ARTEMIS captures artificial emotions in an Agent Knowledge Graph. The Agent Knowledge Graph has

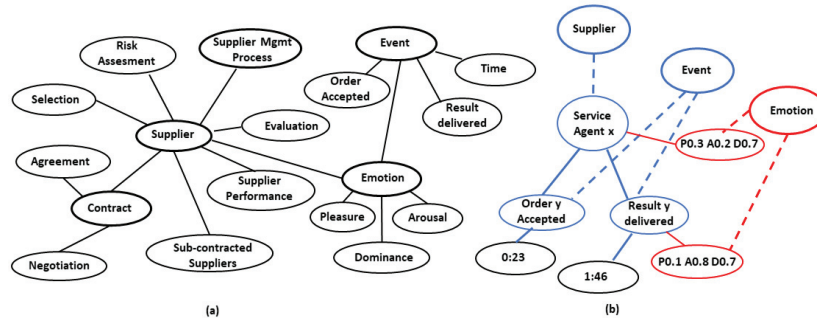


Figure 6 The Agent Knowledge Graph. The Agent Knowledge Graph contains both semantic and episodic information. The semantic part of a Knowledge Graph contains general knowledge about the environment. The Agent Knowledge Graph’s episodic part contains information about specific entities and events that have occurred and the artificial emotions associated with it. (a) The semantic knowledge of the user assistant for the example presented in the motivating example (Section 2) looks like this. The template for this knowledge comes from Graupner [17]. It shows an Agent Knowledge Graph for the process scenario of supplier management. We added the concepts ‘Event’ and ‘Emotion.’ (b) This part of the Agent Knowledge Graph represents information about instances, actual events, and associated artificial emotions. The artificial emotions are associated with the Agent Knowledge Graph with the corresponding interaction events and the causative service providers. The Assistant thus gains an attitude towards the Service Agents over time, which provides useful information for its future selections of cooperation partners.

a semantic and an episodic part. The semantic part of the Agent Knowledge Graph serves to classify information. A protocol of the events that take place represents the basis for the episodic part, such as interactions between the user assistant and the service agents (see the motivating example in Section 2). An Agent Knowledge Graph is an essential part of the ARTEMIS architecture (see Figure 3). Dörner uses a self-defined type of neural network to realize the memory of PSI. For practical reasons, however, we have decided that ARTEMIS will use the established field of knowledge graphs for this purpose.

6.1 Realizing a Semantic Memory in an Agent Knowledge Graph

The semantic memory provides the necessary conceptual information for the assistant about the problem domain. For the motivating example we use knowledge specified by Graupner [17] as basis for the assistant’s semantic part of its Agent Knowledge Graph. Since Graupner’s example is a supplier management system; it is focused around concepts of a “Supplier” and a “Contract”, which are both further related to other concepts (see Figure 6a). ARTEMIS extends this model with the concepts of “Event” and “Emotion”.

6.2 Realizing an Episodic Memory in an Agent Knowledge Graph

While semantic knowledge specifies what an autonomous agent's environment consists of, episodic knowledge describes what is going on in its world. In addition to the abstract semantic knowledge, the user assistant possesses episodic knowledge, such as knowledge about specific service providers, events, or artificial emotions (see Figure 6b). The assistant's interactions with the service providers create episodic knowledge; the user assistant enriches this episodic knowledge with information about its emotions. Over time this leads the user assistant to develop a subjective attitude towards the service providers in its environment. This subjective attitude supports the user assistant in future problem situations and enables selecting appropriate cooperation partners in this complex dynamic environment.

6.3 Capturing Artificial Emotions in the Agent Knowledge Graph

The appraisals of events lead to the creation of artificial emotions. The PAD values represent these artificial emotions, which are initially associated with the events that caused them. Then they are transferred to the service providers involved. The emotion associated with a service agent act as a collection point for all the emotions this service agent has caused over time. Therefore, it is calculated as the weighted sum of positive event emotions minus the weighted sum of negative event emotions. Negative emotions have a greater weight than positive emotions. Note that Kahneman's Prospect Theory [22] supports this imbalance. The theory states, among many other things, that the fear of loss is always more significant than the joy of equal nominal gains.

The Dynamics of Captured Emotions The emotional intensity decreases over time. However, this is not the case if the emotional intensity exceeds a certain intensity threshold. In this case, its intensity does not decrease. This fact applies to emotions associated with events as well as emotions associated with service agents.

7 Experimental Study

We implemented a prototype of the user assistant to assess the performance of ARTEMIS. We aim to answer the following research questions (RQ): **(RQ1)** Can the user assistant generate artificial emotions that are plausible for human test subjects? **(RQ2)** Can captured artificial emotions make the user assistant more efficient?

7.1 Experimental Configuration

To address the proposed research questions, we constructed our experimental configuration as follows:

A Synthetic User Assistant. We implemented a synthetic scenario to evaluate the feasibility and behavior of ARTEMIS. A user assistant is created, which can call 100 service agents. In this scenario, 50 of these service agents are somewhat reliable, and 50 are rather unreliable without the user assistant having any information about them. The user assistant selects its cooperative partners from this pool. It initially selects its cooperation partners at random following a uniform distribution. It can use the artificial emotions generated during the individual interactions and recorded in its knowledge graph during many interactions.

Implementation. We realize the user assistant by a dynamic system based on difference equations; we implemented the system in Python 3.5.3. We modeled the Agent Knowledge Graph as an RDF graph using RDFLib [35]; in order to realize the episodic part of the Agent Knowledge Graph, events are described based on ‘The Event Ontology’ [42].

Experiment. For our experiment, we used two different user assistants based on ARTEMIS: One assistant with the ability to create and store artificial emotion, as described in chapter 5. For the other “emotionless” assistant, we fixed the resulting PAD values as neutral so that the values for pleasure, arousal, and dominance were zero all the time. Before the experiment, the user assistants engaged in a learning phase consisting of 300 test tasks to gather initial knowledge about their environment. During the experiment, both agents had to execute 300 tasks representing user orders and could use their respective knowledge from the learning phase.

User Evaluation: For RQ1, we conducted an evaluation where 30 human test subjects evaluated the user assistant with artificial emotions in the above described scenario. We asked the participants to assess the plausibility of the user assistant’s artificial emotions when fulfilling the user’s order. We presented nine scenarios for each test subject. The basis for the scenarios is the motivational example in Section 2. We asked the test subjects to assess the plausibility of the created artificial emotions within the scenarios. We showed the artificial emotions to the test subjects in both pictorial and textual form.

Evaluation metrics: A way to answer the question of what it means to be ‘better’ in our application scenario (see Section 2) was the following:

Autonomous service agents need a certain amount of time to provide their services. If the user assistant wants to instruct service agents, it must consider their services' expected duration. If an autonomous service agent delivers poor results or no results, the user assistant has a problem, and it is necessary to find and assign another service agent. Finding and assigning another service agent costs additional time. The better the user assistant is in selecting suitable service providers, the less time it takes to fulfill the user's order. The time the user assistant needs to complete an order is a practical measure of the quality of its decisions and its performance. Therefore, concerning RQ2, we measured the two assistants' performance based on ARTEMIS as the elapsed time between the submission of an order to the user assistant and the completion of the order. In our setup, that was measured with the time reported by the Python `time.time()` function.

7.2 Results of the Experimental Study

Results for RQ1. All the users answered the questionnaires independently and evaluated the presented artificial emotions; 270 evaluations were available. Five stated that they would tend to the emotion "indifferent" rather than to the emotion "disdainful" in one of the scenarios in a later optional interview. In nine evaluations, the test persons indicated that they could not decide. In 254 assessments, subjects indicated that they could understand the artificial emotions presented well or very well and imagine having similar emotions in similar situations.

Results for RQ2. Additionally, we evaluated the user assistants' performance in terms of time; we observed the user assistant's behavior with and without artificial emotions. Both user assistants completed 300 runs. As a result, we observed that the effectiveness – in terms of average time – was enhanced by up to 40% whenever the user assistant could fall back on artificial emotions for its decision-making process.

Discussion. As far as we have investigated this, the proposed approach opens up promising research and application fields. These initial results suggest that ARTEMIS's approach works and enables autonomous agents to reach their goals faster. It turns out that remembered artificial emotions help successful agent planning and decision making in complex environments. Also, the experiments' results show that the approach can help make a computer system's decisions more plausible for users. The system can thus make clear its internal situation on which it grounds its decision making. However, further

Table 1 Results of the User Evaluation. We evaluated artificial emotions in a user study; they are represented both as text and as images. In 53.33% of the cases, the users understand the emotions very well, while 40.74% understand them well

User Question	Positive Answers	Percentage Positive Answers %
I fail to understand at all	3	1.11%
I fail to understand	4	1.48%
I cannot decide	9	3.33%
I can understand well	110	40.74%
I can understand very well	144	53.33%

studies considering different scenarios and types of goals, must thoroughly assess the pros and cons of creating and capturing artificial emotions.

8 Conclusions and Future Work

We have tackled the problem of creating and capturing knowledge about artificial emotions. To generate artificial emotions, a suitable model and a system that implements this is required. For this purpose, we have developed the ARTEMIS control system for autonomous agents with artificial emotions. The PSI theory of the cognitive psychologist Dietrich Dörner is the foundation of the ARTEMIS essential components. We added a specific appraisal and an emotional component. In ARTEMIS, event appraisals create artificial emotions. The appraisal pattern described by the emotion psychologist Klaus Scherer is the basis for this. However, Scherer does not provide any information on how to realize this appraisal pattern. ARTEMIS uses Dörner's PSI theory to implement Scherer's appraisal patterns.

For capturing knowledge about artificial emotions, we developed an Agent Knowledge Graph concept to empower autonomous robots and software agents with this knowledge. In addition to knowledge about facts, Agent Knowledge Graphs also represent subjective knowledge of individual autonomous agents. Captured artificial emotions form this subjective knowledge. Artificial emotions are collected together with other information (e.g., point in time) about events in Agent Knowledge Graphs. As time goes by, the captured artificial emotions form a subjective world view of the agents. This subjective world view enhances the agent's ability to plan and decide successfully in complex dynamic environments. An essential aspect of the artificial emotions created by ARTEMIS is that they have a meaning. This meaning can be derived as follows:

The basis of the emotions generated is appraisal processes evaluating the impact of external events on the agent. The appraisals influence the need processes, which represent the agent's internal state. These need processes generate values for the parameters "Pleasure," "Arousal," and "Dominance" that define a so-called PAD space. In our version of the PAD space, each point represents one of the eight proposed emotions with a specific intensity. Finally, according to Gaerdenfors [16], the PAD space has the potential to transmit meaning to the created emotions. That makes ARTEMIS emotions an abstract and expressive representation of the agent's evaluation of events concerning its internal state.

We empirically investigated ARTEMIS's behavior in a synthetic scenario in which a user assistant selected suitable cooperation partners from a pool of 100 service agents. In three hundred interactions, the user assistant developed an emotional attitude toward many of these service providers. We have evaluated the feasibility of the artificial emotions the assistant created by a group of thirty human test subjects. The test subjects confirmed that most of the user assistant's artificial emotions were plausible to them. Furthermore, we measured the user assistant's execution time in settings with and without remembered artificial emotions. The evaluation results show that a user assistant can reach their objective on average in 40% less time than the configuration without remembered artificial emotions. The observed results reveal the potential of ARTEMIS. Nevertheless, we recognize that this formalism is still in an initial phase and that further studies are required to provide a general approach that can represent artificial emotions in various scenarios.

For the near future, we are planning: (1) To further improve the performance of the system based on more experiments, and (2) to enhance the communication between ARTEMIS and its human users. Furthermore, we aim to tackle the problem of tracking down ARTEMIS decisions. Thus, empowering ARTEMIS to explain the behavior during the achievement of a user assistant's objectives is part of our future work.

References

- [1] J.R. Anderson and C. Lebiere. *The atomic components of thought*. Lawrence Erlbaum Associates Publishers, 1998.
- [2] E. André, M. Klesen, P. Gebhard, S. Allen, and T. Rist. Integrating Models of Personality and Emotions into Lifelike Characters. In *Affective Interactions*, LNAI 1814, 2000.

- [3] D. Antos and A. Pfeffer. Using Emotions to Enhance Decision-Making. *Proceedings of IJCAI*, 2011.
- [4] C. Becker, S. Kopp, and I. Wachsmuth. Simulating the Emotion Dynamics of a Multimodal Conversational Agent. In *Affective Dialogue Systems*, LNCS, 3068, 2004.
- [5] C. Breazeal. *Designing Sociable Robots*. MIT Press, 2002.
- [6] D. Cañamero. Modeling motivations and emotions as a basis for intelligent behavior. In *Proceedings of the First International Symposium on Autonomous Agents*. doi 10.1145/267658.267688, 1997.
- [7] F. Detje. *Handeln erklären*. DUV Deutscher Universitätsverlag, doi 10.1007/978-3-663-08224-8, Bamberg, 1999.
- [8] D. Dörner and T. Stäudel. Emotion und Kognition. In K.R. Scherer (Eds.), *Psychologie der Emotion*, Göttingen, 1990.
- [9] D. Dörner, H. Schaub, and F. Detje. Das Leben von PSI. Über das Zusammenspiel von Kognition, Emotion und Motivation – oder: Eine einfache Theorie komplexer Verhaltensweisen. In R. von Lüde, D. Moldt, and R. Valk (Eds.), *Sozionik aktuell*, 2, Informatik Universität Hamburg, 2001.
- [10] D. Dörner. *Die Mechanik des Seelenwagens – Eine neuronale Theorie der Handlungsregulation*. (The mechanics of the soul car – A neuronal theory of action regulation). Huber, Bern, 2002.
- [11] D. Dörner. The Mathematics of Emotion. In *The Logic of Cognitive Systems*, Fifth International Conference on Cognitive Modelling, Universitätsverlag Bamberg, 2003.
- [12] D. Dörner and C.D. Güss. PSI: A computational architecture of cognition, motivation, and emotion. In *Review of General Psychology*, 17, 3, 297–317, 2013.
- [13] J. Dominique, D. Fensel, J. Davies, R. González-Cabero, and C. Pedrinaci. The service web: a web of billions of services. In G. Tselentis, J. Domingue, A. Galis, A. Gavras, D. Hausheer, S. Krco, V. Lotz, and T. Zahariadis (Eds.), *Towards a Future Internet: a European Research Perspective*, 203–216, IOS Press, Amsterdam, 2009.
- [14] P. Ekman and D. Cortavro. What is meant by calling emotions basic. In *Emotion review*, 3, 4, 364–370, 2011
- [15] S.C. Gadanho. Learning Behavior-Selection by Emotions and Cognition in a Multi-Goal Robot Task. In *Journal of Machine Learning Research*, 4, 385–412, 2003.

- [16] P. Gaerdenfors. *The Geometry of Meaning: Semantics based on Conceptual Spaces*. The MIT Press, Cambridge, Massachusetts, London, 2014.
- [17] S. Graupner, H.R.M. Netzad, and S. Singhal. Making Processes from Best Practice Frameworks Actionable. In *Proceedings of 13th Enterprise Distributed Object Computing Conference Workshops*, 2009.
- [18] N.M. Hakak, M. Mohd, M. Kirmani, and M. Mohd. Emotion Analysis: A Survey. In *Proceedings of International Conference on Computer, Communications and Electronics (Comptelix)*, Manipal University Jaipur, Malaviya, 2017.
- [19] C. Hoffmann and M.E. Vidal. Creating and Capturing Artificial Emotions in Autonomous Robots and Software Agents. In Bielikova M., Mikkonen T., Pautasso C. (eds), *Web Engineering, ICWE 2020 Lecture Notes in Computer Science*, 12128, Springer, Cham, 2020.
- [20] E. Hudlicka. What are we modelling when we model emotions? In *AAAI spring symposium: emotion, personality, and behavior*, 8, 2008.
- [21] H. Jiang and M.N. Huhns. EBDI: an architecture for emotional agents. In *Proceedings of 6th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2007)*, Honolulu, Hawaii, USA, 2007.
- [22] D. Kahneman and A. Tversky. Prospect Theory: An Analysis of Decision under Risk. In *Econometrica* 47.2 (1979): 263–292, 1979.
- [23] Z. Kowalczyk and M. Czubenko. Computational Approaches to Modelling Artificial Emotion – An Overview of the Proposed Solutions. *Front. Robot. AI* 3-21, 2016.
- [24] J.E. Laird, A. Newell, and P.S. Rosenbloom. Soar: An architecture for general intelligence In *Artificial intelligence*, 33.1, 1–64. 1987.
- [25] J. Lin, M. Spraragen, and M. Zyda. Computational models of emotion and cognition. In *Advances in Cognitive Systems*, 2012.
- [26] R.P. Marinier, J.E. Laird, and R.L. Lewis. A computational unification of cognitive behavior and emotion. In *Cognitive Systems Research*, 10, 1, 48–69, 2009.
- [27] S.C Marcella and J. Gratch. EMA: A process model of appraisal dynamics. In *Cognitive Systems Research* 10, 2009.
- [28] S. Marsella, J. Gratch, and P. Petta. Computational Models of Emotion. In K.R. Scherer, T. Bänziger, and E. Roesch (eds.), *A Blueprint for Affective Computing: A Sourcebook and Manual*, OXFORD UNIV PR, 2010.

- [29] A. Mehrabian. Framework for a comprehensive description and measurement of emotional states. In *Genetic, Social, and General Psychology Monographs*, 121, 339–361, 1995.
- [30] A. Mehrabian. Pleasure-Arousal-Dominance: A General Framework for Describing and Measuring Individual Differences in Temperament. In *Current Psychology: Developmental Learning Personality*, Social Winter, 14, 4, 261–292, 1996.
- [31] A. Moors, P.C. Ellsworth, K.R. Scheerer, and N.H. Frijda. Appraisal theories of emotion: State of the art and future development. In *Emotion Review*, 2, 119–124, 2013
- [32] A. Newell. Unified theories of cognition. Harvard University Press, 1994.
- [33] A. Ortony, G.L. Clore, and A. Collins. The Cognitive Structure of Emotions. NY: Cambridge, 1988.
- [34] A. Rao and M. Georgeff. Modelling rational agents within a BDI-architecture. In *Proceedings of the International Conference on Principles of Knowledge Representation and Planning*, Morgan Kaufmann publishers, San Mateo, California, 1991.
- [35] RDFLib Homepage. <https://rdflib.readthedocs.io/en/stable/>. Last accessed 2020-10-28.
- [36] S.J. Russell, and P. Norvig. Artificial intelligence: a modern approach. Prentice Hall, 2009.
- [37] M.A. Salichs and M. Malfaz. A New Approach to Modeling Emotions and their Use on Decision-Making System for Artificial Agents. In *IEEE Transactions on Affective Computing*, 3, 1, 2012.
- [38] K.R. Scherer. On the nature and function of emotions: A component process approach. In *Approaches to emotion*, 2293, 1984.
- [39] K.R. Scherer. The component process model: Architecture for a comprehensive computational model of emergent emotion. In K.R. Scherer, T. Bänziger, and E.B. Roesch (eds.), *Blueprint for Affective Computing: A Sourcebook*, 47–70, Oxford University Press, 2010.
- [40] C.A. Smith and P.C. Ellsworth. Patterns of Cognitive Appraisal in Emotion. In *Journal of Personality and Social Psychology*, 48, 4, 813–838, 1985.
- [41] C.A. Smith and R.S. Lazarus. Emotion and Adaptation. In *Handbook of Personality: Theory and Research*, 1990.
- [42] The Event Ontology. <http://motools.sourceforge.net/event/event.html> Last accessed: 2020-10-28.

- [43] J.D. Velásquez. When robots weep: Emotional memories and decision-making. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence*, 1999.
- [44] Verein Berliner Kaufleute und Industrieller. Leitsätze ehrbarer Kaufleute. <https://www.vbki.de/leitsaetze> last accessed: 2020-10-25.
- [45] Wikipedia. Honourable Merchant. https://en.wikipedia.org/wiki/Honourable_Merchant last accessed: 2020-10-25.
- [46] M. Wooldridge, *An Introduction to Multi Agent Systems*. John Wiley and Sons, United Kingdom, 2009.

Biographies



Claus Hoffmann is a computer scientist and has been working in research since 1984. He studied computer science with a focus on artificial intelligence. His scientific interests are decision models for complex situations, user/assistant agents, and autonomous agents with artificial motivations and emotions.



Pascal Linden is a computer science student in the master's program of the University of Bonn, Germany. The current focus of his studies is on the

exploration of different aspects of intelligent systems. Besides his studies he works as a student assistant at the Fraunhofer Institute for Intelligent Analysis and Information Systems on semantic web technologies.



Maria-Esther Vidal is head of the Scientific Data Management research group at TIB and a member of the L3S research center at Leibniz Universität Hannover. She is also a full professor (on leave) at the Universidad Simón Bolívar (USB) Venezuela. Her research interests include data and knowledge management, knowledge representation, and the semantic web. She has published more than 170 papers on Semantic Web, databases, bioinformatics, and artificial intelligence.

